Homology modeling and molecular dynamics simulations of MUC1-9/H-2K^b complex suggest novel binding interactions

Athanassios Stavrakoudis,[1] Ioannis G. Tsoulos,[2] Katalin Uray,[3]
Ferenc Hudecz[3,4] and Vasso Apostolopoulos[5]

[1] Department of Economics, University of Ioannina, Ioannina, Greece
[2] Department of Communications, Informatics and Management, Technological Educational Institute of Epirus, Arta, Greece
[3] Research Group of Peptide Chemistry, Hungarian Academy of Sciences, Budapest, Hungary
[4] Institute of Chemistry, Eötvös Loránd University, Budapest, Hungary
[5] Immunology and Vaccine Laboratory, Centre for Immunology, The Macfarlane Burnet Institute for Medical Research and Public Health, Melbourne, Australia

Correspondence:
Athanassios Stavrakoudis, Department of Economics, University of Ioannina, GR-451 10, Ioannina, Greece, tel: +30 265 100 5935, fax: +30 265 100 5092, email: `astavrak@cc.uoi.gr`, web: `http://stavrakoudis.econ.uoi.gr`

Vasso Apostolopoulos, Immunology and Vaccine Laboratory, Centre for Immunology, The Macfarlane Burnet Institute for Medical Research and Public Health, Melbourne, Australia tel: +61 3928 22111, fax: +61 3928 22100, email: `vasso@burnet.edu.au`

Abstract

Human MUC1 is over-expressed on human adenocarcinomas and has been used as a target for immunotherapy studies. The 9-mer MUC1-9 peptide has been identified as one of the peptides which binds to murine MHC class I H-2K^b The structure of MUC1-9 in complex with H-2K^b has been modeled and simulated with classical molecular dynamics, based on the x-ray structure of the SEV9 peptide/H-2K^b complex. Two independent trajectories with the solvated complex (10 ns in length) were produced. Approximately 12 hydrogen bonds were identified during both trajectories to contribute to peptide/MHC complexation, as well as 1-2 water mediated hydrogen bonds. Stability of complex was also confirmed by buried surface area analysis, although the corresponding values were about 20% lower than those of the original x-ray structure. Interestingly, a bulged conformation of the peptide's central region, partially characterized as a β-turn, was found exposed form the binding groove. In addition, P1 and P9 residues remained bound in the A and F binding pockets, even though there was suggestion that P9 was more flexible. The complex lacked the numerous water mediated hydrogen bonds that were present in the reference peptide x-ray structure. Moreover, local displacements of residues Asp4, Thr5 and Pro9 resulted in loss of some key interactions with the MHC molecule. This might explain the reduced affinity of the MUC1-9 peptide, relatively to SEV9, for the MHC class I H-2K^b.

Keywords: Class I MHC; H-2K^b; homology modeling; molecular dynamics; MUC1; tumor

1 Introduction

Major histocompatibility complex (MHC) proteins bind small peptide fragments derived from pathogenic proteins and form peptide/MHC (pMHC) complexes (Raghavan et al., 2008). MHC proteins are divided into two classes: class I (MHC-I) and class II (MHC-II). The MHC-I consists of a polymorphic transmembrane heavy chain and β2-microglobulin, which are non-covalently associated (Zhang et al., 1998). The proteolysis of intracellular proteins by the proteasome produces the majority of peptides suitable for MHC-I binding. In most cases, peptides of 8-10 residues in length are found in the binding groove of MHC-I.

After the first crystal structures of pMHC complexes were available, (Bjorkman et al., 1987; Fremont et al., 1992; Matsamura et al., 1992) it was suggested that peptides bound to MHC-1 with a canonical extended strucure. MHC class I residues that form the binding groove are responsible for the specificity of the peptide selection. Six (out of 8-10) residues of the peptide sequence are accomodated within the A-F binding pockets of the MHC-I protein (Saper et al., 1991). Residues that do not participate directly in binding are believed to interact with the TCR.

Human mucin, MUC1, is a membrane-bound glycoprotein, expressed on the surface of epithelial cells. It is often overproduced and/or underglycosylated in adenocarcinomas (breast, ovary, colon, lung, kidney, etc) and is present in the serum of cancer patients. MUC1 is immunogenic in mice and in humans, with both humoral and cellular immune responses being induced by MUC1-based vaccine constructs (Tang et al., 2008b,a). MUC1 mucin partly consists of a variable number of tandem repeats region of the consensus sequence ^{1}PDTRPAPGSTAPPAHGVTSA^{20} which is repeated 40-80 times (Gendler et al., 1988). The majority of anti-MUC1 antibodies recognize sequences within the SA^{1}PDTRPAP^{7} region (Price et al., 1991; Xing et al., 1991, 1992; Burchell et al., 1989). The SAPDTRPAP (MUC1-9) 9-mer peptide was also found to be presented by MHC-1 H-2K$^{\mathrm{b}}$ and to be immunogenic (Apostolopoulos et al., 1997). MUC1-9 binds with low affinity to H-2K^{b} (Apostolopoulos

et al., 1997) via a noncanonical mode and it was suggested that the C-terminus of the peptide looped out of the peptide binding groove (Apostolopoulos et al., 1998; Apostolopoulos and Lazoura, 2004).

Computer simulation of molecular dynamics is a well established method for studying several aspects of biomolecular structure and function (Hansson et al., 2002; Karplus, 2003; Aksimentiev et al., 2008; Tantar et al., 2008). In recent years such computational approaches have been increasingly incorporated in drug design (Galeazzi, 2009), in immunological reasearch (Morikis and Lambris, 2004; Mallik and Morikis, 2006; Stavrakoudis, 2010) and also to peptide/MHC complexes (Omasits et al., 2008; Knapp et al., 2009). Moreover, biomolecular modeling can complement experimental studies (van Gunsteren et al., 2008) and can elucidate dynamics of immunological synapse (Wan et al., 2008), allows to study the dynamics of a peptide bound to antibody (Tatsis et al., 2009; Stavrakoudis, 2009b), could be used to model disulphide peptide complexed proteins (i.e. C8γ (Stavrakoudis, 2009a)) or even more excitingly to help in clinical decision making (Sadiq et al., 2008).

Modeling of the MUC1-9 peptide with both murine and human MHC class I, H-2K^{b} and HLA-A2 respectively have been previously performed (Apostolopoulos et al., 1998), based on a simulated annealing protocol and high temperature molecular dynamics (Chelvanayagam et al., 1996). That work was a considerable progress in our knowledge of peptide/MHC interactions in the MUC1-9 case and provided a possible structural explanation of the antibody binding of MUC1 peptides presented by the MHC molecules. However, modern progress in computational biophysics, accompanied with the big enhancement of available computer power, can be utilized to further improve the computer-generated model of the MUC1-9 peptide complexed the the MHC class I H-2K^{b}.

Here, we present a homology modeling and molecular dynamics approach of MUC1-9 (SAPDTRPAP) in complex with MHC class I H-2K^{b}. Since the initial conformation was modeled rather than taken from an x-ray structure, we chose

to perform two indepent simulation runs, to obtain more robust results. Long-run dynamics, inclusion of the whole MHC molecule and explicit representation of solvent have been utilized in order to more accurate picture the MUC1-9 structure and interactions with the MHC molecule. Such approach has been suggested to give more reliable results in MD investigations (van Gunsteren et al., 2008; Omasits et al., 2008). Our results suggest that this was a beneficial approach in the current study, and has given insights into the peptide binding mode of the MUC1-9.

2 Methods

Initial coordinates for the SEV9/MHC complex were downloaded from Protein Data Bank (Berman et al., 2002), access code: 1kpv.

The original peptide from sendai-virus, FAPGNYPAL was mutated to SAPDTR-PAP, whilst MHC molecule remained untouched. The SEV9 peptide was selected from other canditates due to its homology with the MUC1-9 peptide. Pro residue homology in positions P3 and P7 was also crucial for selection. Since the backbone dihedral angle ϕ of Pro residue is restrained, it is preferable to choose a peptide that has the same residue in these positions. Ideally, it would be perfect to also have alignment for position P9, however there was no such option. Topology and force field parameters for all atoms were assigned from the CHARMM22-CMAP parameter set (Mackerell et al., 2004; MacKerell et al., 2004). It has been noted that addition of cross terms with CMAP potential improves the system parametrization and helps to avoid undesired backbone helical transitions (Buck et al., 2006; Stavrakoudis, 2008).

Hydrogen atoms were added with the VMD program (Humphrey et al., 1996) and its autopsf utility. Protonation status of Histidine side chains were determined with the REDUCE program (Word et al., 1999). The peptide/MHC complex was centered in a rectagular box with dimensions $95.7 \times 88.3 \times 102.9 \text{Å}^3$. The box was filled

with TIP3P water molecules and neutralized with the addition of 26 Na^+ and 20 Cl^- ions respectively, to approximate a 0.1 mM ion concetration. Crystallographic water molecules (345) were also included in the model. The final system contained 24429 water molecules. Total number of atoms of the entire system were 80598.

Non-bonded van der waals interactions were gradually turned off at a distance between 12 and 14 Å (Yonetani, 2006). Long range electrostatics were calculated with the PME method (Darden et al., 1993). Non-bonded forces and PME electrostatics were computed every second step. Pair list was updated every 10 steps. Bonds to hydrogen atoms were constrained with the SHAKE method allowing a 2 fs time step for integration. The system was initially subjected to energy minimization with 5 000 steps. The temperature of the system was then gradually increased to 310 K, with Langevin dynamics using the NVT ensemble, during a period of 3 000 steps, by stepwise reassignment of velocities every 500 steps. The simulation was continued at 310 K for 100 000 steps (200 ps). During minimization and equilibration phases, protein backbone atoms (N, C^{α}, C', O) and oxygen atoms of crystallographic waters were restrained to their initial positions with a force constant of 50 kcalmol $^{-1}$Å^{-2}. The system was equilibrated for further 200 ps with the force constant reduced to 5 kcalmol^{-1}Å^{-2}. Finally, 400 ps of NVT simulation at 310 K was performed with total elimination of the positional restraints. The simulation was passed to the productive phase, by applying constant pressure with the Langevin piston method (Feller et al., 1995). Velocities were re-initialized and two independent trajectories were produced (trA and trB). Pressure was maintained at 1 atm and temperature at 310 K. Results are based to a period of 10 ns of this isothermal-isobaric (NPT) runs. Shapshots were saved to disk at 1 ps interval for structural analysis.

The initial structure of the SEV9/MHC complex (PDB code 1kpv) were also simulated under identical conditions for comparative analysis (tr0 trajectory).

Trajectory analysis was performed with Eucb (Tsoulos and Stavrakoudis, 2009) and Carma (Glykos, 2006) software packages. Secondary structure analysis was

performed with STRIDE (Frishman and Argos, 1995). Circular data statistics (dihedral angles, etc) were calculated with appropriate corrections (Agostinelli, 2009). Structural figures were prepared with PyMOL (www.pymol.org).

2.1 Burried surface area calculation

Calculation of buried surface area (BSA) was performed with the NACCESS program (http://www.bioinf.manchester.ac.uk/naccess/), based on the formula:

$$BSA = S_p + S_a - S_c \tag{1}$$

thus as the difference of the surface accessible area of the complex (S_c) from the sum of the of surface accessible areas of the peptide (S_p) and MHC molecule (S_a) respectively.

β-turn classifications were based on geometrical characteristics of the backbone conformation (Hutchinson and Thornton, 1994). Initially, a β-turn was accepted if $d(C_i^\alpha - C_{i+3}^\alpha) \leq 7\mathring{A}$ and $|\alpha(C_i^\alpha - C_{i+1}^\alpha - C_{i+2}^\alpha - C_{i+3}^\alpha)| < 90\circ$, where d is the distance and a is the dihedral angle between the corresponding atoms. Further classification of the β-turn was based on hydrogen bond patterns and backbone dihedral values of the $i+1$ and $i+2$ residues.

In order to indentify isolated (from the bulk) water molecules in the peptide/MHC interface the instantaneous water coordination number (N_c) approach (Petrone and Garcia, 2004). This method counts the water oxygen atoms within a range (typically 3.5 Å) of any water oxygen atom, which is actually the first hydration shell. The N_c can be found between 0 and 15, depending on the local structure of water. In the bulk water this number is always greater than 3, while in the protein interior is 0 to 2. This implies that a water molecule has no other water neighbours and it is inside the protein interior. The N_c is measured for all the MD trajectory and isolated water molecules are indentified if the N_c value is small for a prolonged period of

time. In the current study, a search of water molecule with $N_c \leq 1$ for at least 70% of the MD time has been performed.

2.2 MM-PBSA calculation of $\Delta G_{binding}$

The binding free energy of the association of two molecules (A+B→AB) can be estimated, according to the MM-PBSA approach (Kollman et al., 2000; Wan et al., 2005), as:

$$\Delta G_{binding} = G^{AB} - G^{A} - G^{B}, \tag{2}$$

where:

$$\Delta G^{i} = \langle E_{MM} \rangle + \langle G_{solv} \rangle - TS. \tag{3}$$

In the above equations, $\langle . \rangle$ denotes average value for a a set of snapshots algon a molecular dynamics trajectory, while E_{MM} is the molecular mechanics energy of the i^{th} molecule in the gas phase, namely the sum of f internal bonded energy (comprising bond, angle and dihedral terms), van der Waals and electrostatic interactions. G_{solv} is the solvation free energy of the i^{th} molecule. This term can be estimated as the sum of the electrostatic solvation free energy calculated by the Poisson–Boltzmann equation and the non-polar solvation free energy calculated from the SASA.

Hence, the binding free energy is:

$$\Delta G^{i} = \langle \Delta E_{MM} \rangle + \langle \Delta G_{solv} \rangle - TS. \tag{4}$$

The average properties can be computed directly from the MD trajectory snapshots. In the current study, the last 5 ns were used, assuming that equilibrium was reached after the first 5 ns of the simulation. 5000 structures were utilized for the SASA and E_{MM} calculations, while 50 structures (one every 100 frames) were used for the calculation of the G^{elec}_{solv} with the APBS (Baker et al., 2001; Dolinsky et al., 2004) software.

3 Results and Discussion

3.1 RMSF and RMSD analysis

Root mean square fluctuations (RMSF) of the C^{α} atoms of the MHC and peptide chains, as well as the time evolution of the root mean square deviation (RMSD) of the backbone atoms (N, C^{α}, C') of the MHC and peptide chains, during both MD trajectories, trA and trB respectively, are shown in Figure 1.

In both trA and trB cases, RMSF profiles of chains A and B from the MHC molecule were almost identical, which indicates the robustness of the study. RMSF values were between 0.5 and 2.0 Å, which is quite common in similar MD studies of protein complexes around equilibrium. Similarly, RMSD time series were also very similar for chain A and B, with only a small exception of the trA trajectory: RMSD values escaped from stationarity around 4ns in trA, and a small peak of RMSD 0.28 Å was observed. In general, both trajectories were quite stable, Figure 1. Time series of RMSD fluctuated around 1.5–2.0 Å for chain A and around 1.0 Å for chain B. If we take into consideration the simulation temperature (310 K) these values are considered small, indicating the stability of the complex. Moreover, there is strong evidence that the MHC molecule did not undertake significant conformational changes upon mutation of the peptide residues (Fremont et al., 1992; Matsamura et al., 1992). This is in accordance with other X-ray studies of the H-$2K^b$ MHC class-I molecule with different nonamer peptides in the binding groove. These observations corroborate our hypothesis that homology modeling coupled with molecular dynamics simulations produces a reliable model of the MUC1-9/H-$2K^b$ complex.

Peptide's RMSF values of C^{α} atoms showed an interesting differentation between trA and trB trajectories. While values of 0.5-1.0 Å were recorded for residues 1-7 in both cases, trA trajectory showed increased values of 1.5 and 2.0 Å for residues 8 and 9 respectively. In trB trajectory, RMSF remained close to 1.0 Å for all

residues. Values in the order of 2.0 Å are still considered relatively small, however, the differentation is notable. Since this fact was observed in only one of the two trajectories, it could be considereded as a relative random effect of the simulation. On the other hand, it definitely indicates that the peptide binding to the MHC groove is not so tight at the C-terminal region, as previously has been suggested (Apostolopoulos et al., 1998; Apostolopoulos and Lazoura, 2004).

Peptide's RMSD time series of backbone atoms were very similar in both cases. RMSD values ranged between 0.99 and 2.1 Å and averaged at 1.46 (0.16) Å for trA, whilst the RMSD values ranged between 1.04 and 1.69 Å and averaged at 1.36 (0.09) for trB case. There is only a minor difference between these two profiles: trB trajectory showed slightly smaller values with smaller standard deviation of the time series. This is possibly due to increased fluctuation at the C-terminal end in trA . However, as it was previously noted for chains A and B of the MHC molecule, RMSD profiles of the peptide corroborate the stability of the trajectories and the validity of the proposed model.

3.2 Peptide backbone dynamics

Backbone conformations play an important role in peptide/MHC binding (Barinaga, 1992; Matsamura et al., 1992). Here we present a detailed analysis of the peptide's backbone conformation.

Figure 2 displays the distribution (Ramachandran map) of the backbone dihedral angles ϕ, ψ of peptide residues in the region 2-8. It is evident that, for most of the residues, the backbone dihedrals show very similar distributions in the trA and trB trajectories. The only exception comes from the **Ala8** residue. As it has been noted, the C-terminal residues showed increased mobility (higher RMSF values), and this is very well reflected in the distribution of its backbone dihedral angles.

The initial values of -61° and 150° of ϕ, ψ angles of **Ala2** were well conserved during both MD trajectories. Percentage of dihedral angles from both trA and trB

trajectories within 30° of the initial values were found 98% and 80% for ϕ, ψ angles respectively.

Pro3's backbone ϕ, ψ dihedral angles was -60° and 146° respectively in the initial structure. As it was expected, the fluctuation of ϕ was found rather small, and over 60% of the frames were found within 15° of the initial value (>99% if 30° bin is taken into consideration). Backbone ψ angle also showed minimal fluctuation and more than 80% of the frames in both trA and trB trajectories were found within 30° of the initial value.

Asp4's backbone ϕ, ψ dihedral angles were -120° and 153° respectively in the initial structure. Contrary to the **Pro3** case, **Asp4** residue experienced a significant move to its backbone ϕ dihedral angle. Time series of this angle fluctuated between -30° and -122° and averaged at -69°(11°). Only 35% of the trA frames and 55% of the trB frames remained within 30° of the initial value. Similarly, backbone ψ angle averaged at -35°(12°). Thus **Asp4** residue showed (in total) an approximately 100° move in backbone dihedral angles. It could be considered that **Asp4** represents a first differentation between the crystal structure of the reference peptide and the MUC1-9 peptide studied here.

Thr5's backbone ϕ, ψ dihedral angles were 74° and 48° respectively in the initial structure of the SEV9 peptide. A positive ϕ angle, although abnormal in other cases, is not uncommon in peptide's conformation of other peptide/MHC complexes. For example ϕ angle of residue Ser5 was found to be 60° in SRDHSRTPM (YEA9) peptide (Apostolopoulos et al., 2002). During both trA and trB trajectories, the sign of backbone ϕ dihedral angle of residue **Thr5** changed quickly and the residue adopted backbone ϕ angles close to –150° (Figure 2). Time series of **Thr5**'s ϕ angle averaged at -151°(22°) in both trA and trB trajectories. Negative values of ϕ at position 5 have also been observed in other crystal structures of peptide/MHC H-2K^b complexes. For example, in the SSYRRPVGI peptide from influenza A virus, the ϕ angle of Arg5 was found to be -67° (PDB access code 1wbz) (Meijers et al.,

2005). The identical results obtained in both trajectories underline the robustness of the found values for **Thr5**'s ϕ angle. Backbone dihedral ψ of **Thr5** averaged at 162°(65°) and 160°(49°) in trA and trB trajectories respectively. Average values are approximately 115° different from the initial value.

Arg6's backbone ϕ, ψ dihedral angles was -59° and 107° respectively in the initial structure. Similarly to **Thr5**, backbone dihedral angles were altered during MD trajectories. Average values of ϕ angle were found to be -129°(14°) and -128°(13°) in trA and trB trajectories respectively. Average values of ψ angle were found to be 153°(17°) and 151°(12°) in trA and trB trajectories respectively. Only 47% of trajectories frames in trA and 30% in the trB retained backbone dihedrals within 30° of the initial values.

Pro7's backbone ϕ, ψ dihedral angles was -57° and 144° respectively in the initial structure. Average values of ϕ angle were found to be -49°(13°) and -53°(13°) in trA and trB trajectories respectively. Average values of ψ angle were found to be 138°(21°) and 143°(19°) in trA and trB trajectories respectively. After three continuous residues that escaped the initial conformation, **Pro7** retained mostly its initial structure.

Ala8's backbone ϕ, ψ dihedral angles were -65° and 145° respectively in the initial structure. **Ala8**'s backbone ϕ angle averaged at -119°(28°) and -118°(19°) during trA and trB MD trajectories respectively. As it is indicated by the higher standard deviation value, and it is also seen in Figure 2, values of ϕ backbone dihedral showed significant more dispersion during trA trajectory than in trB. This is in accordance with the higher RMSF value observed for **Ala8** in the trA trajetory. Backbone dihedral angle ψ was found to be similar to its initial values. Average values of ψ angle were found to be 140°(21°) and 130°(15°) in trA and trB trajectories respectively.

Pro9's ϕ dihedral angle remained close to -70° (as it is expected from the proline's cyclic structure). The original (from the x-ray structure, Leu9) anlge was

-70.9°. Thus, there was no significant backbone difference in this part of the peptide.

Hairpin and β-turn structures in peptides bound to MHC molecules have been identified in case of MHC class II molecules (Zavala-Ruiz et al., 2004). However, this happens to the peptide's region that is outside of the binding group. In the current study, we have identified a very interesting case of β-turn in the central region of the peptide, covering residues **Pro3** to **Arg6**. This sequence has been found in β-turn conformation for 50 and 77% of the simulation time, in the trA and trB trajectories respectively. We did not recorded any intra-peptide hydrogen bond stabilizing this β-turn. Table 1 lists the values of backbone dihedral angles as calculated for the central residues of the β-turn, **Asp4** and **Thr5** respectively. Both trA and trB trajectories showed very close values of backbone ϕ and ψ dihedrals. These values differ from the initial values found in the crystal structure of the SEV9 peptide. However, the common finding from the two independent trajectories (trA and trB) corroborate the suggestion that a β-turn around the **Asp4-Thr5** region exists, at least partially.

3.3 Interactions between the peptide and the MHC

The binding mode of nonamer peptides with the H-2K^{b} MHC class I molecule has been investigated in the past. There are numerous studies in the literature (Matsamura et al., 1992; Fremont et al., 1992; Apostolopoulos et al., 2002; Meijers et al., 2005; D. H. Fremont and E. A. Stura and M. Matsumura and P. A. Peterson and I. A. Wilson, 1995) addressing the principles of peptide anchoring to MHC's binding groove. It is generally assumed that H-2K^{b} has six binding pockets, A to F, that accomodate residues P1,P2,P3,P6,P7 and P9 of nonamer peptides (Matsamura et al., 1992; Saper et al., 1991). Residues P4 and P5 do not make direct contacts with the MHC molecule and protrude towards the solvent, hence their side chains are available for interaction with the TCR. The charge groups of N- and C-terminal

residues make strong interactions with the MHC binding clefts (pockets A and F respectively).

A general view of the peptide/MHC binding motif in shown in Figure 3, whilst the peptide's orientation inside the MHC's binding is depicted at Figure 4.

Peptide's **Ser1** (P1) was found to form two stable hydrogen bonds with the MHC molecule. Its backbone atoms N and O were found in hydrogen bond state with side chains of Glu63A and Tyr159A respectively. These hydrogen bonds were conserved, in both trA and trB trajectories, for approximately 91 to 95% of the simulation time (Table 2). The distance between Ser1:N and Glu63A side chain oxygen atoms, in the initial structure, were found 4.6 and 5.8 Å for $O^{\epsilon 1}$ and $O^{\epsilon 2}$ respectively, which indicates that this strong (charged) hydrogen bond between the N-terminal group of the peptide and the side chain of Glu63A was formed during the modeling proccess and was not present in the initial structure. Indeed, Glu63A's side chain (atom $O^{\epsilon 1}$) actually was to form a hydrogen bond with Ala2:N atom, in the structure of the original peptide (Matsamura et al., 1992). The hydrogen bond between Ser1:O and Tyr159A:O^{η}, on the other hand, was well formed in the initial structure (distance 2.67 Å) and very well conserved in both MD trajectories (Table 2). Another hydrogen bond interaction between Ser1 and the MHC molecule was present between the side chains of Ser1 and Tyr7A (or Tyr171A for short periods), for approximately 95% of the simulation time. This is very interesting, since no side-chain interactions have been observed in the x-ray structure of SEV9 peptide (Fremont et al., 1992). Thus, overall two to three hydrogen bonds contributed to peptide's binding. These results corroborate the importance of this binding pocket in the peptide/MHC binding process.

Side chain of Glu63A (pocket B) accepted hydrogen bond from **Ala2** Nitrogen atom (position P2). This interaction was conserved for 93.5% (trA) or 98.6% (trB) of the simulation time, and it was well formed in the initial structure (the distance between Ala2:N and Glu63A:$O^{\delta 1}$ was found 2.9 Å). This finding underlines the

importance of the Glu63A residue, since its negatively charged side chain formed two stable hydrogen bonds with the peptide's backbone amide groups. Side chain of Lys66A was found in hydrogen bond state with Ala2:O atom for over 90% of the simulation time. The corresponding distance between Lys66A:N^{ζ} and Ala2:O atoms in the initial structure was found 2.7 Å, indicating the existense of the hydrogen bond. Moreover, side chains of Tyr7A and Tyr45A made hydrophobic contacts with **Ala2**'s aliphatic side chain. The above analysis is for the **Ala2** interactions is almost identical with the x-ray structure of the SEV9 peptide (Fremont et al., 1992), indicating the fact the preservation of the **Ala2** residue in position P2 (binding pocket B) contributed to the retaing of the same peptide/MHC interactions.

Pro3 (P3) made important hydrophobic interactions with Tyr159A's side chain. Average distance of their side chain centers were found 4.0 Å(0.6) or 4.2 Å(0.6) during trA or trB MD trajectories respectively. For approximately 25% of the time, the two side chains were found in parallel orientation forming a stacking interaction. It is noted that Tyr159A's side chain donated a hydrogen bond to Ser1:O, hence this MHC residue is considered to contribute significantly to peptide's binding. The original hydrogen bond between Pro's backbone oxygen atom and Asn70A's side chain was found to be relatively weak during trA and trB MD trajectories: 12.7 and 27.2% of the frames respectively satisfied the hydrogen bond criteria.

Central residues **Asp4** and **Thr5** did not show any significant interactions with the MHC's residues. Only **Asp4**'s side chain was found hydrogen bonded to Arg62A's side chain for limited period of simulation time, $\approx 15\%$. Both residues were exposed outside of the binding groove.

Binding pocket C plays an important role in peptide recognition by MHC H-$2K^{b}$ molecules (Molano et al., 1998; D. H. Fremont and E. A. Stura and M. Matsumura and P. A. Peterson and I. A. Wilson, 1995; Huard et al., 1997). Peptide's residue **Arg6** side chain at position P6, was found to form a strong hydrogen bond with Glu24A side chain. Actually, these side chains remained hydrogen bonded the entire

time time in both trA and trB MD trajectories. On the other side, there was no backbone interaction with the MHC molecule. However, the ability of the MHC molecule to bind different peptide sequences, since the original peptide has Tyr in this position (Apostolopoulos et al., 2002), which is a canonical residue at this position for MHC binding. Tyr6 (SEV9 peptide) to Arg6 (MUC1-9 peptide) mutation led to some loss of hydrophobic interactions between peptide and MHC molecule, a fact that might explain the reduced binding affinity of the MUC1-9 peptide, relative to SEV9 peptide. However, the **Arg6** remained inside the canonical C-pocket, unlike the **Arg6** residue in YEA9 peptide (SRDNSRIPM) which utilized the non-canonical E binding pocket (Apostolopoulos et al., 2002).

Residue **Pro7**, at peptide's P7 position, had a weak backbone hydrogen bond with Tyr117A's side chain. Occurence was found 28% in trA and only 7% in trB trajectories respectively. Given the fact that in crystal structures of peptides bound in the H-2K^b molecule, no such hydrogen bond exist (Table 2), the result is not so suprising. However, significant hydrophobic interactions with Trp147A and Trp133A side chains were found to contribute in peptide/MHC interactions. For example, side chain distances between **Pro7** and Trp147A varied between 3 and 5 Å and averaged at 3.6 Å (0.2). To a lesser degree, Leu156A and Tyr116A also made hydrophobic contacts with side the chain of **Pro7**.

Position P8 was occupied by **Ala8**. The backbone carbonyl group of this residue was found to be in hydrogen bond state with Trp147A's side chain. This is a well expected interaction, as it has been found in the crystal structure of the original peptide. A relatively weak hydrogen bond was also formed for part of trA trajectory, between Ala8:N and Glu152A:Oϵ2. The corresponding distance in the initial structure was found to be 5.8 Å.

Finally, residue **Pro9** at position P9 (binding pocket F). The C-terminal carboxyl group was found to form two hydrogen bonds (Table 2) with Thr143A and Lys146A side chains, for almost all of the simulation time, in both trA and trB tra-

jecoties. The same interactions were also present in the x-ray sructure that served as initial point for these calculations. However, the lack of amide hydrogen in proline's structure resulted to the abolishment of a backbone hydrogen bond between peptide and the MHC molecule. Thus, the Leu to Pro (SEV9 to MUC1-9 peptide) mutation resulted in a small shift of the position of this residue. These subtle changes in peptide's conformation have been shown (Hoare et al., 2008) to affect drastically the MHC recognition and might explain to some extent the reduced affinity of the MUC1-9 peptide when bound to class I H-2K^{b}. **Pro9**'s side chain also made hydrophobic contacts with Val76A, Leu81A and Trp147A side chains. For at least 90% of the simulation time, a pair of side chain heavy atoms from these residues were in close contact (distance less than 4.5 Å) with a side chain heavy atom from **Pro9**. These hydrophobic interactions further stabilized the peptide/MHC interactions, and along with the hydrogen bonds strengthen the anchoring role of **Pro9**.

Overall, as it can be seen from Figure 6, there were approximately 12 hydrogen bonds between the peptide and the MHC molecule, during both MD trajectories. This number approximates very well the number of the reported (Fremont et al., 1992) hydrogen bonds (11) between the peptide SEV9 and the MHC molecule.

3.4 Buried Surface Area

Buried surface area (BSA) is a good indicator of the binding of a ligand into a protein (Olsson et al., 2008). Figure 7 shows the time evolution of BSA between the peptide and the MHC molecule. BSA flucuated between 666.9 and 1005.6 Å and averaged at 848.7(47.5) Å in the trA trajectory. In the trB case, BSA values were found between 656.7 and 999.6 Å with mean value of 824.9(52.9) Å. As it is can be drawn from the graphical representaion of BSA time evolution, and from basic statistical analysis, both trajectories showed similar profiles for the calculated BSA of peptide/MHC interface. The difference of approximately 25 Å^2 (3%) in the mean values is very small and could be considered to be within expected error. In

a recent experimental re-investigation of BSA of protein x-ray structures (Novotny et al., 2007) it was suggested that differences from 50 to 100 $Å^2$ in BSA values were expected as a measurement error rather than actual difference in BSA. These findings corroborate our statement that the peptide/MHC complex was stable and that the fluctuations in BSA time series are normal.

The BSA value in the x-ray structure of the SEV9 peptide was 1076 $Å^2$, while the BSA value in the MUC1-9/MHC complex after restrained energy minimization was found to be 937.5 $Å^2$. The loss of approximately 140 $Å^2$ can be attributed to minor conformational changes that occured during MD run in order for the mutated peptide to adapt to the binding groove of the MHC molecule. Considered, however, that trajectories were obtained in 310 K, thus the spontaneous thermal moving of the atoms resulted in somewhat reduced BSA values.

Thus, the difference of more than 140 $Å^2$ in the BSA of the peptide/MHC interface, in the SEV9 and MUC1-9 cases, is another indication of the lower binding affinity that the MUC1-9 has to the H-2K^b molecule, relatively to the SEV9 peptide.

4 Concluding remarks

Homology modeling and molecular dynamics simulations have been used to assess the structure of the SAPDTRPAP/H-2K^b complex. Results presented here indicate that a stable complex is formed, based on the analysis of two MD trajectories.

MHC binding pockects A and F interacted closely with the N- and C-terminus of the peptide which played an important role in stabilizing the complex. The Buried Surface Area of the peptide/H-2K^b interface remained constant during the simulation indicating the stability of the complex and its similarity to the initial peptide/MHC complex.

Replacement of Leu with Pro at P9 position did not affect significantly the MHC's binding of the peptide. The C-terminal carboxyl group was found to form

stable hydrogen bonds with the MHC molecule, and the non-polar side chain of Pro residue made a number of close contacts with hydrophobic residues of the MHC's F binding pocket. However, the peptide showed relatively increased mobility in the C-terminal region, that may affet the strength of the MHC binding.

A main difference between MUC1-9's simulated structure and SEV9's x-ray structure was the ϕ angle of **Thr5**. A significant transition from +74° to ≈ -150° occured. Since it is well known that backbone conformation plays a very important role in peptide/MHC recognition (Barinaga, 1992), it is expected that this conformational transition would alter the MHC's binding affinity for the peptide, most possibly downwards. Moreover, MHC H-2K^{b} molecules prefer hydrophobic residues at position P6 (for nonamer peptides), even though MUC1-9 has Arg in this place.

This has resulted in a notable alteration of the backabone conformation of the central part of peptide and the enhancement of the exposure of the Asp4-Thr5 region outside of the MHC's binding groove. For a considerable amount of simulation time this bulged region adopted a β-turn conformation, however without the presence of the characteristic hydrogen bond. This had not been noted in previous modeling studies (Apostolopoulos et al., 1998) and provides a new framework for the peptide/MHC interactions.

Inclusion of explicit water molecules in the current study helped a lot to clarify the role of the solvent in peptide/MHC interactions. Water mediated hydrogen were found only sparingly and although existed, a clear contribution to the binding procces can not be attributed to this kind of interaction.

Leu to Pro muation at position P9 resulted in slight movement of this residue within the F binding pocket. However, this fact, along with the loss of a hydrogen bond interaction of the Leu amide hydrogen might be enough reason for observing the reduced affinity of the MUC1-9 peptide to H-2K^{b} binding.

All of the above observations reflected well in the reduction of the BSA between the peptide and the MHC molecule, where a loss of 140 $\AA^2$ has been measured.

Finaly, it seems that while the MUC1-9 peptide forms stable complex with the H-$2K^b$ molecule, it is clear that certain structural reorganization occured and resulted in reduced binding affinity.

Acknowledgements

NAMD parallel execution have been performed at the Research Center of Sientific Simulations (RCSS) of the University of Ioannina. The open source community is gratefully acknowledged for proving all the necessary tools (Linux, NAMD, GNU, etc) that made this work possible.

References

Agostinelli, C., 2009. Circular Statistics with R.
URL `http://cran.r-project.org/web/packages/circular/index.html`

Aksimentiev, A., Brunner, R., Cohen, J., Comer, J., Cruz-Chu, E., Hardy, D., Rajan, A., Shih, A., Sigalov, G., Yin, Y., Schulten, K., 2008. Computer modeling in biotechnology: a partner in development. Methods Mol Biol 474, 181–234.

Apostolopoulos, V., G Chelvanayagam, P. X. X., McKenzie, I. F., 1998. Anti-MUC1 antibodies react directly with MUC1 peptides presented by class I H2 and HLA molecules. J Immunol 161, 767–775.

Apostolopoulos, V., Haurum, J. S., McKenzie, I. F., 1997. MUC1 peptide epitopes associated with five different H-2 class I molecules. Eur J Immunol 27, 2579–2587.

Apostolopoulos, V., Lazoura, E., 2004. Noncanonical peptides in complex with MHC class I. Expert Rev Vaccines 3, 151–162.

Apostolopoulos, V., Yu, M., Corper, A. L., Li, W., McKenzie, I. F. C., Teyton, L., Wilson, I. A., Plebanski, M., 2002. Crystal structure of a non-canonical high affinity peptide complexed with MHC class I: a novel use of alternative anchors. J Mol Biol 318, 1307–1316.

Baker, N. A., Sept, D., Joseph, S., Holst, M., McCammon, J. A., 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad USA 98, 10037–10041.

Barinaga, M., 1992. Getting some "backbone": how MHC binds peptides. Science 257, 880–881.

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D.,

Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J. D., Zardecki, C., 2002. The Protein Data Bank. Acta Crystall Sec D: Biol Crystall 58, 899–907.

Bjorkman, P. J., Saper, M. A., Samraoui, B., Strominger, W. S., Wiley, D. C., 1987. Structure of the human class I histocombatibility antigen, HLA-A2. Nature 329, 506–512.

Buck, M., Bouguet-Bonnet, S., Pastor, R., MacKerell, A., 2006. Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. Biophysical Journal 90, 36–38.

Burchell, J., Taylor-Papadimitriou, J., Boshell, M., Gendler, S., Duhig, T., 1989. A short sequence, within the amino acid tandem repeat of a cancer-associated mucin, contains immunodominant epitopes. Int J Cancer 44, 691–696.

Chelvanayagam, G., Jakobsen, B., Gao, X., Easteal, S., 1996. Structural comparison of major histocompatibility complex class i molecules and homology modelling of five distinct human leukocyte antigen-a alleles. Protein Engineering 9, 1151–1164.

D. H. Fremont and E. A. Stura and M. Matsumura and P. A. Peterson and I. A. Wilson, 1995. Crystal structure of an H–2K^{b}–ovalbumin peptide complex reveals the interplay of primary and secondary anchor positions in the major histocompatibility complex binding groove. Proc Natl Acad Sci USA 92, 2479–2483.

Darden, T., York, D., Pedersen, L., 1993. Particle mesh ewald: An $N \log(N)$ method for ewald sums in large systems. J Chem Phys 98, 10089–1092.

Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., Baker, N. A., 2004. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. Nucl Acids Res 32, W665–W667.

Feller, S. E., Zhang, Y., Pastor, R. W., Brooks, B. R., 1995. Constant pressure molecular dynamics simulation: The Langevin piston method. J Chem Phys B 103, 4613–4621.

Fremont, D. H., Matsamura, M., Stura, E. A., Peterson, P. A., Wilson, I. A., 1992. Crystal structures of two viral peptides in complex with murine MHC class I H-2K^{b}. Science 257, 919–927.

Frishman, D., Argos, P., 1995. Knowledge-based protein secondary structure assignment. Proteins: Structure, Function and Genetics 23, 566–579.

Galeazzi, R., 2009. Molecular Dynamics as a Tool in Rational Drug Design: Current Status and Some Major Applications. Current Computer-Aided Drug Design 5, 225–240.

Gendler, S., Taylor-Papadimitriou, J., Duhig, T., Rothbard, J., Burchell, J., 1988. A highly immunogenic region of a human polymorphic epithelial mucin expressed by carcinomas is made up of tandem repeats. J Biol Chem 263, 12820–12823.

Glykos, N. M., 2006. Software news and updates. Carma: a molecular dynamics analysis program. J Comput Chem 27, 1765–8.

Hansson, T., Oostenbrink, C., van Gunsteren, W. F., 2002. Molecular dynamics simulations. Curr Opin Struct Biol 12, 190–196.

Hoare, H., Sullivan, L., Clements, C., Ely, L., Beddoe, T., Henderson, K., Lin, J., Reid, H., Brooks, A., Rossjohn, J., 2008. Subtle Changes in Peptide Conformation Profoundly Affect Recognition of the Non-Classical MHC Class I Molecule HLA-E by the CD94–NKG2 Natural Killer Cell Receptors. J Mol Biol 377, 1297–1303.

Huard, R., Dyall, R., Nikolić-Žugić, J., 1997. The critical role of a solvent-exposed residue of an MHC class I-restricted peptide in MHC-peptide binding. International Immunology 9, 1701–1707.

Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: Visual Molecular Dynamics. J Mol Graph 14, 33–38.

Hutchinson, E. G., Thornton, J. M., 1994. A revised set of potentials for beta-turn formation in proteins. Protein Sci 3, 2207–22016.

Karplus, M., 2003. Molecular dynamics of biological macromolecules: a brief history and perspective. Biopolymers 68, 350–358.

Knapp, B., Omasits, U., Frantal, S., Schreiner, W., 2009. A critical cross-validation of high throughput structural binding prediction methods for pMHC. J Comp Aid Mol Des 23, 301–307.

Kollman, P., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., et al., 2000. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33, 889–897.

MacKerell, A. D., Feig, M., Brooks, C. L., 2004. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J Comput Chem 25, 1400–1415.

Mackerell, A. D., Feig, M., Brooks, C. L., 2004. Improved treatment of the protein backbone in empirical force fields. J Am Chem Soc 126, 698–699.

Mallik, B., Morikis, D., 2006. Applications of Molecular Dynamics Simulations in Immunology: A Useful Computational Method in Aiding Vaccine Design. Curr Proteom 3, 259–270.

Matsamura, M., Fremont, D. H., Peterson, P. A., Wilson, I. A., 1992. Emerging principles for the recognition of peptide antigens by MHC class I molecules. Science 257, 927–934.

Meijers, R., Lai, C. C., Yang, Y., Liu, J. H., Zhong, W., Wang, J. H., Reinherz, E. L., 2005. Crystal Structures of Murine MHC Class I H-2 D^b and K^b Molecules in Complex with CTL Epitopes from influenza A Virus: Implications for TCR Repertoire Selection and Immunodominance. J Mol Biol 345, 1099–1110.

Molano, A., Erdjument-Bromage, H., Fremont, D., Messaoudi, I., Tempst, P., Nikolic-Zugic, J., 1998. Peptide selection by an MHC H-2Kb class I molecule devoid of the central anchor ("C") pocket. The Journal of Immunology 160, 2815–2823.

Morikis, D., Lambris, J. D., 2004. Physical methods for structure, dynamics and binding in immunological research. Trends Immunol 25, 700–707.

Novotny, M., Seibert, M., Kleywegt, G. J., 2007. On the precision of calculated solvent-accessible surface areas. Acta Crystallographica Section D 63, 270–274.

Olsson, T. S. G., Williams, M. A., Pitt, W. R., Ladbury, J. E., 2008. The thermodynamics of protein–ligand interaction and solvation: Insights for ligand design. J Mol Biol 384, 1002–1007.

Omasits, U., Knapp, B., Neumann, M., Steinhauser, O., Stockinger, H., Kobler, R., Schreiner, W., 2008. Analysis of key parameters for molecular dynamics of pMHC molecules. Molecular Simulation 34, 781–793.

Petrone, P., Garcia, A., 2004. MHC–peptide binding is assisted by bound water molecules. J Mol Biol 338, 419–435.

Price, M., F., H., O'Sullivan, C., Baldwin, R., Edwards, P., Tendler, S., 1991. Immunological and structural features of the protein core of human polymorphic epithelial mucin. Mol Immunology 27, 795–802.

Raghavan, M., Del Cid, N., Rizvi, S., Peters, L., 2008. MHC class I assembly: out and about. Trends in Immunology 29, 436–443.

Sadiq, S. K., Mazzeo, M. D., Zasada, S. J., Manos, S., Stoica, I., Gale, C. V., Watson, S. J., Kellam, P., Brew, S., Coveney, P. V., 2008. Patient-specific simulation as a basis for clinical decision-making. Phil Trans R Society A 366, 3199–3219.

Saper, M. A., Bjorkman, P. J., Wiley, D. C., 1991. Refined Structure of the Human Histocompatibility Antigen HLA-A2 at 2.6 Å Resolution. J Mol Biol 219, 277–319.

Stavrakoudis, A., 2008. Molecular dynamics simulations of an apoliprotein derived peptide. Chem Phys Lett 461, 294–299.

Stavrakoudis, A., 2009a. A disulfide linked model of the complement protein C8γ complexed with C8α indel peptide. J Mol Model 15, 165–171.

Stavrakoudis, A., 2009b. Computational modeling and molecular dynamics simulations of a cyclic peptide mimotope of the CD52 antigen complexed with CAMPATH-1H antibody. Mol Sim 36, 127–137.

Stavrakoudis, A., 2010. Conformational Flexibility in Designing Peptides for Immunology: The Molecular Dynamics Approach. Curr Comput-Aided Drug Design 6, in press. URL `http://www.ncbi.nlm.nih.gov/pubmed/20412039`

Tang, C.-K., Katsara, M., Apostolopoulos, V., 2008a. Strategies used for muc1 immunotherapy: human clinical studies. Expert Rev Vaccines 7, 963–975.

Tang, C.-K., Katsara, M., Apostolopoulos, V., 2008b. Strategies used for muc1 immunotherapy: preclinical studies. Expert Rev Vaccines 7, 951–962.

Tantar, A. A., Conilleau, S., Parent, B., Melab, N., Brillet, L., Roy, S., Talbi, E.-L., Horvath, D., 2008. Docking and Biomolecular Simulations on Computer Grids: Status and Trends. Cur Comput Aid Drug Des 4, 235–249.

Tatsis, V. A., Tsoulos, I. G., Stavrakoudis, A., 2009. Molecular Dynamics Simulations of the TSSPSAD Peptide Antigen in Free and Bound with CAMPATH-1H Fab Antibody States: The Importance of the β-Turn Conformation. Int J Pept Res Ther 15, 1–9.

Tsoulos, I. G., Stavrakoudis, A., 2009. Eucb: a C++ program for trajectory analysis, http://stavrakoudis.econ.uoi.gr/eucb.

van Gunsteren, W. F., Dolenc, J., Mark, A. E., 2008. Molecular simulation as an aid to experimentalists. Curr Opin Struct Biol 18, 149–53.

Wan, S., Coveney, P., Flower, D., 2005. Peptide recognition by the T cell receptor: comparison of binding free energies from thermodynamic integration, Poisson–Boltzmann and linear interaction energy approximations. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 363, 2037.

Wan, S., Flower, D. R., Coveney, P. V., 2008. Toward an atomistic understanding of the immune synapse: large-scale molecular dynamics simulation of a membrane-embedded TCR-pMHC-CD4 complex. Mol Immunol 45, 1221–1230.

Word, J. M., Lovell, S. C., Richardson, J. S., Richardson, D. C., 1999. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J Mol Biol 285, 1735–1747.

Xing, P.-X., Prenzoska, J., McKenzie, I., 1992. Epitope mapping of anti-breast and anti-ovarian mucin monoclonal antibodies. Mol. Immunology 29, 641–650.

Xing, P.-X., Reynolds, K., Pietersz, G., McKenzie, I., 1991. Effect of varitions in peptide sequence on anti-human milk fat globule membrane antibody reactions. Immunology 72, 304–311.

Yonetani, Y., 2006. Liquid water simulation: a critical examination of cutoff length. J Chem Phys 124, 204501.

Zavala-Ruiz, Z., Strug, I., Walker, B. D., Norris, P. J., Stern, L. J., 2004. A hairpin turn in a II MHC-bound peptide orients outside the binding groove for T cell recognition. Proc Nat Acad Sci USA 101, 13279–13284.

Zhang, C., Anderson, A., DeLisi, C., 1998. Structural principles that govern the peptide-binding motifs of class I MHC molecules. J Mol Biol 281, 929–947.

Tables

Table 1 Backbone dihedral angles in the region Asp4-Thr5 of the MUC1-9 peptide, where a β-turn was found in the MUC1-9 peptide. Corresponding region of the SEV9 peptide is Gly4-Asn5. Column PDB lists the corresponding values from the crystal structure of the SEV9 peptide, with the residues Gly and Asn at positions 4 and 5 respectively. Averages values (and variances in parentheses) is fiven from trajectories tr0, trA and trB.

Dihedral	**PDB**	**tr0**	**trA**	**trB**
ϕ_4	-119.9	-138.3 (4.9)	-67.8 (1.1)	-69.4 (1.2)
ψ_4	153.2	-173.3 (3.6)	-34.8 (1.3)	-35.4 (1.2)
ϕ_5	74.4	58.8 (0.9)	-151.3 (2.1)	-151.4 (2.2)
ψ_5	48.2	49.2 (1.1)	161.7 (3.5)	160.0 (3.9)

Table 2 Hydrogen bond interactions between the SEV9 and MUC1-9 peptides and the H-2K^{b} molecule. Percentage of frames is given, from trajectories tr0, trA and trB, that met the geometrical criteria for hydrogen bond interaction. Distance between donor-acceptor atoms are taken from the initial structure (PDB column).

Donor	**Acceptor**	**PDB** (Å)	**tr0** (%)	**trA** (%)	**trB** (%)
Phe_{1P}:N	Tyr_{59A}:O^{η}	4.13	32.4		
Phe_{1P}:N	Glu_{63A}:$O^{\epsilon 1,2}$	4.61	95.6		
Ser_{1P}:N	Glu_{63A}:$O^{\epsilon 1,2}$	5.83		93.2	94.1
Tyr_{159A}:O^{η}	Ser_{1P}:O	2.67	59.6	91.4	94.8
Tyr_{59A}:O^{η}	Ser_{1P}:O^{γ}	5.77		77.7	81.9
Ser_{1P}:O^{γ}	Tyr_{7A}:O^{η}	5.82		93.2	97.9
Ser_{1P}:O^{γ}	Tyr_{171A}:O^{η}	5.00			20.6
Ala_{2P}:N	Glu_{63A}:$O^{\epsilon 1,2}$	2.90	90.1	93.5	98.6
Lys_{66A}:N^{ζ}	Ala_{2P}:O	2.74	76.6	97.3	90.0
Asn_{70A}:N^{δ}	Pro_{3P}:O	3.63	63.4	12.7	27.2
Arg_{62A}:$N^{\eta 2}$	Asp_{4P}:$O^{\delta 1,2}$	6.78		19.0	
Arg_{6P}:$N^{\eta 1,2}$	Glu_{24A}:$O^{\epsilon 1,2}$	5.21		92.4	97.2
Tyr_{116A}:O^{η}	Pro_{7P}:O	4.11		28.2	7.2
Ala_{8P}:N	Glu_{152}:$O^{\epsilon 1,2}$	5.87		38.5	5.5
Trp_{147A}:N^{ϵ}	Ala_{8P}:O	2.86	11.7	70.4	98.5
Leu_{9P}:N	Asp_{77A}:$O^{\delta 1,2}$	3.02	82.4		
Lys_{146A}:N^{ζ}	Leu_{9P}:$O^{\tau 2}$	3.00	95.2		
Tyr_{84A}:O^{η}	Leu_{9P}:$O^{\tau 2}$	2.84	39.3		
Thr_{143A}:O^{γ}	Pro_{9P}:$O^{\tau 1,2}$	2.68		94.1	92.7
Lys_{146A}:N^{ζ}	Pro_{9P}:$O^{\tau 1,2}$	3.00		98.6	98.7

Figures

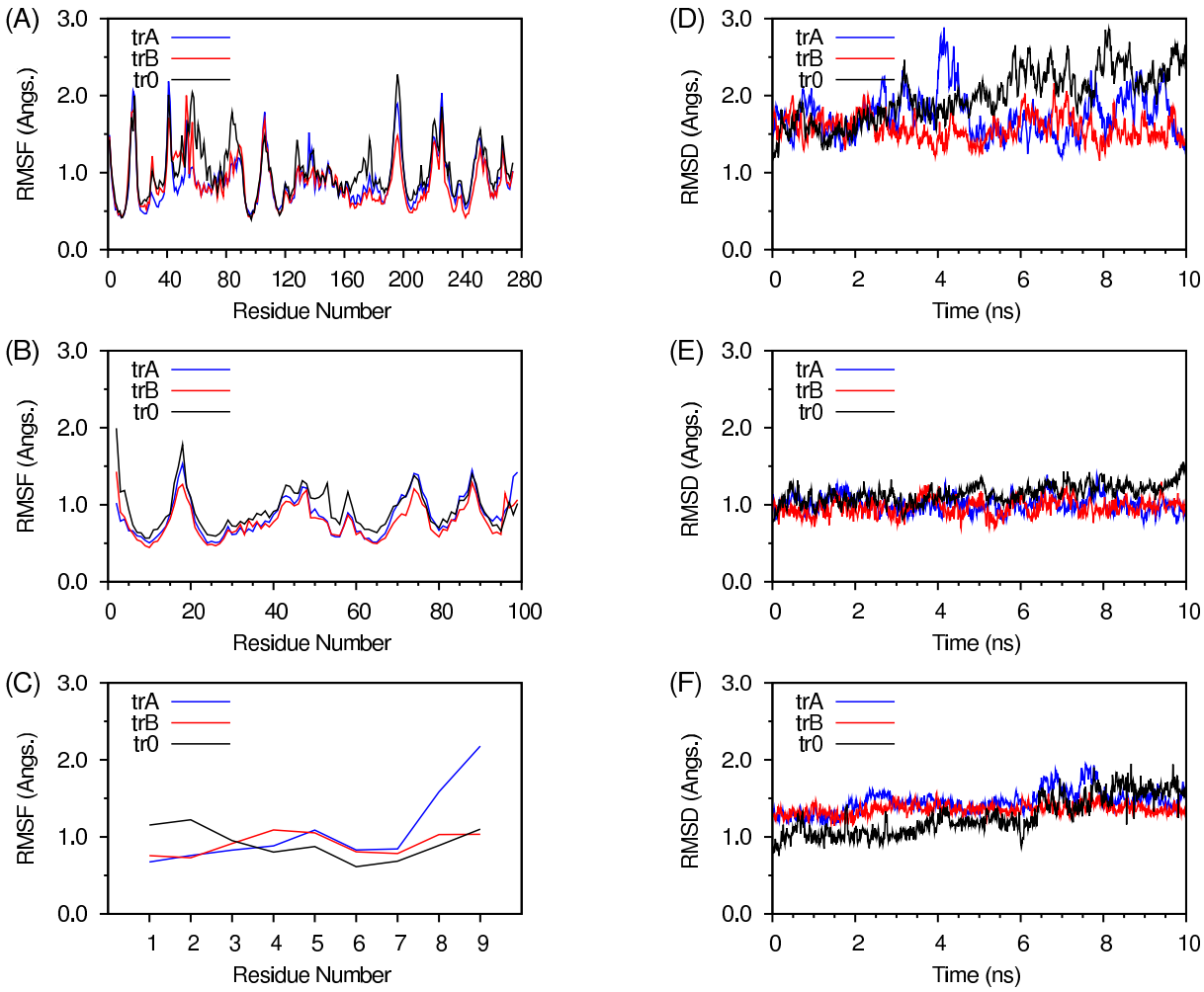

Figure 1: Root mean square fluctuation (left column) of C^{α} atoms and root mean square deviation (right column) time series of backbone atoms (N, C^{α}, C') of the pMHC complex after fitting the corresponding atom positions from MD trajectory to initial (X-ray) coordinates. Results from different trajectories (tr0, trA and trB) are indicated with different line colors. A) RMSF of MHC chain A, B) RMSF of MHC chain B, C) RMSF of MHC chain P (peptide), D) RMSD of MHC chain A, E) RMSD of MHC chain B and F) RMSD of MHC chain P (peptide).

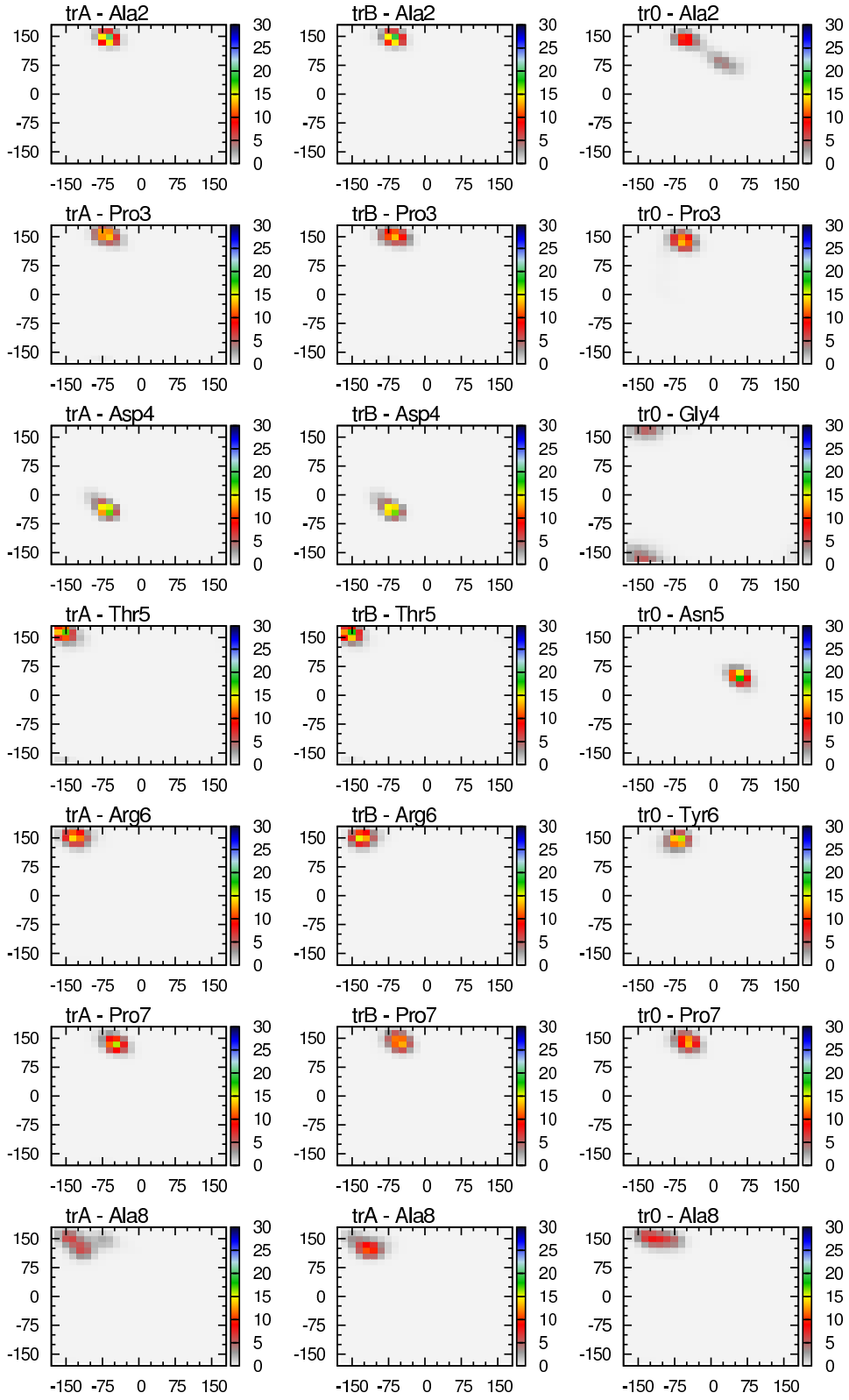

Figure 2: Ramachandran plot of backbone dihedral angles of the peptide. Horizontal axis is for ϕ and vertical axis is for ψ angle respectively. The plots represent propability density maps, z-axis is the percentage of frames found within 10° dihedral angle bin. The adjacent colour bar is used to identify regions of low (grey) versus high (blue) populations.

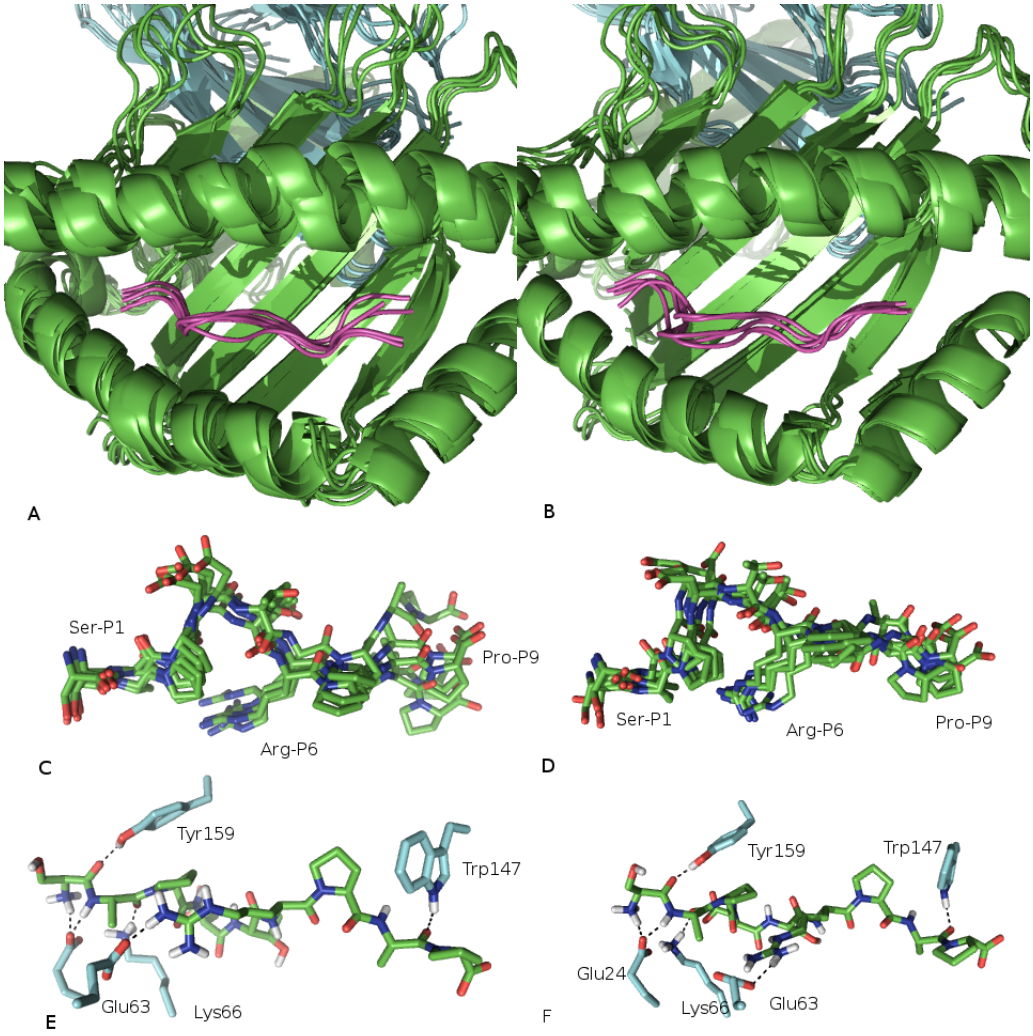

Figure 3: A) Ribbon representation of five selective structures of the MUC1-9/H-2K^{b} complex (one frame every 2 ns) from trA trajectory, B) Ribbon representation of five selective structres of the MUC1-9/H-2K^{b} complex (one frame every 2 ns) from trB trajectory, C) Stick representation of the peptide bound in the MHC groove from trA trajectory, D) Stick representation of the peptide bound in the MHC groove from trB trajectory, E) Important hydrogen bond interactions between the pepetide and MHC molecule in the trA trajectory and F) Important hydrogen bond interactions between the peptide and MHC molecule in the trB trajectory. Hydrogens were omitted from stick represantations. Structures have been fitted to the first frame using the backbone atoms.

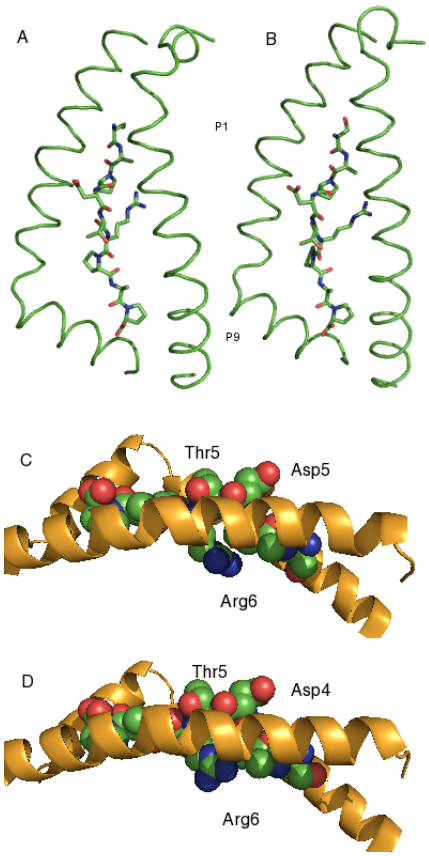

Figure 4: Peptide's (sticks) orientation in MHC (ribbons) binding groove in trA (A) and trB (B) trajectories. Exposure to the solvent of the region **Asp4-Thr5**, while **Arg6** side chain orientates towards the beta-sheet floor, in the interior of the binding groove of pocket C, in trA (C) and trB (D) trajectories.

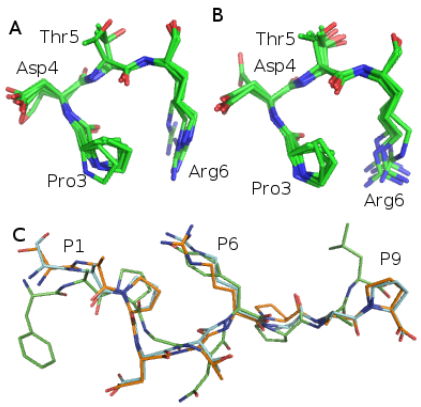

Figure 5: A) and B) Backbone overlay of the Pro3-Arg6 region of the peptide from the trA and trB trajectories respectively. This fragment has been found in β-turn conformation for considerable amount of time. C) Backbone superimposition of SEV9 peptide (green) from the X-ray structure with representative structures from trA (cyan) and trB (orange) trajectories. The differentation of backbone conformation at fragment **Asp4-Thr5** is well seen. Side chains of residues 2, 3, 6 and 7 share common orientation towards the MHC binding groove. Interestingly, conformations of residues at positions 1 and 9 deviate from the original structure.

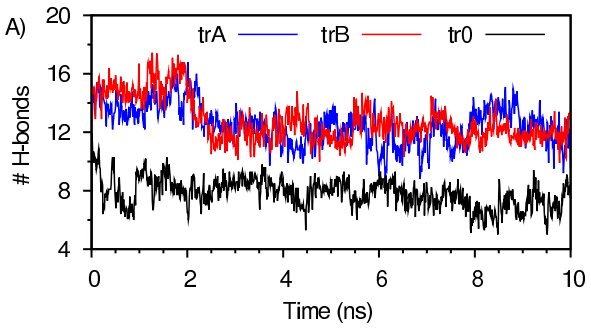

Figure 6: A) Total number of hydrogen bonds between the peptide and the MHC molecule, as evolved over simulation time. Data were averaged every 10 ps. B) Total number of water mediated hydrogen bonds between the pepetide and MHC molecule, as evolved over simulation time. Data were taken every 10 ps.

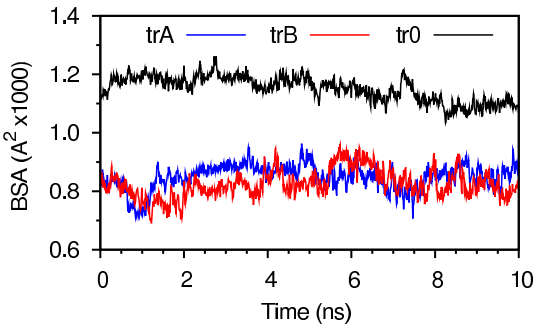

Figure 7: Time series of apolar buried surface area (BSA) between the peptide and the MHC molecule in tr0, trA and trB trajectories.